

## SISTEM KOMPUTASI BLACKBOX UNTUK OPTIMASI PENGKOREKSIAN MULTI TIPE DAN TEKNIK SKORSING SOAL OBYEKTIF

Lilik Hidayati<sup>1</sup>, Ripai<sup>2</sup>

<sup>1,2</sup> FMIPA UNW Mataram

<sup>1</sup>liliknyagatan@gmail.com, <sup>2</sup>ripainasir@yahoo.co.id

### Abstrak

Tes sebagai alat seleksi maupun evaluasi harus bermutu agar dapat menghasilkan nilai yang obyektif dan akurat sehingga penilaian yang diberikan kepada siswa bersifat adil. Secara empiris, tes disebut bermutu apabila memenuhi persyaratan/sifat (1) Option pengecoh berfungsi, (2) memiliki tingkat kesukaran yang ideal, (3) dapat membedakan kelompok atas dan bawah, (4) Valid dalam mengukur kemampuan siswa dan (5) reliable/konsisten dalam memberikan penilaian. Kesulitan guru dalam melakukan analisis mutu tes karena harus dilakukan secara kuantitatif sesuai deskripsi parameter matematis dan statistik persyaratan mutu dari respon jawaban siswa. Kondisi ini lebih sulit dilakukan ketika kebijaksanaan pemerintah dalam ujian nasional menggunakan model soal dengan multi tipe, dimana konstruksi indikator soal secara acak, menyulitkan guru untuk melakukan pengoreksian manual dengan berbagai tipe kunci jawaban. Lebih lanjut guru juga harus dapat melakukan penyusunan kembali option jawaban siswa menurut kesamaan indikator. Hal ini diperlukan karena analisis mutu tes dan harus didasarkan atas butir dengan konten isi mengukur indikator yang sama. Penelitian dilakukan guna mengembangkan sistem komputasi *BlackBox* yang dapat melakukan uji mutu tes secara simultan dari model multi tipe tes pilihan ganda. Hasil pengembangan yang telah dilakukan menggunakan *software* MATLAB membuahkan 10 (sepuluh) *m-function* yang dapat saling berintraksi untuk melakukan komputasi uji mutu tes secara simultan. (1) *m-function* susun untuk menyusun option jawaban siswa menurut butir indikator tes, (2) *m-function* koreksi, untuk mengoreksi jawaban siswa, (3) *m-function* kualitas pengecoh untuk analisis berfungsinya option pengecoh, (4) *m-function* tingkat kesukaran untuk analisis tingkat kesukaran butir tes, (5) *m-function* dayabeda untuk analisis butir soal yang dapat membedakan kelompok atas dan bawah, (6) *m-function* poinbiserial, untuk mengukur kesahihan (validitas) butir soal dalam mengukur kemampuan siswa, (7) *m-function* KR20 untuk mengukur konsistensi (reliabilitas) butir soal dalam memberikan penilaian kepada siswa, (8) *m-function* seleksi butir untuk membuang butir soal yang tidak memenuhi syarat mutu, (9) *m-function* skor dan nilai untuk memberikan skor dan penilaian kepada siswa berdasarkan butir soal yang bermutu dan (10) *m-function* uji mutu tes untuk melakukan analisis uji mutu tes secara simultan. Hasil pengembangan *m-function* tersebut akan dikembangkan lagi menjadi sistem komputasi berbasis *Grafik User Interface* (GUI) dan aplikasi *project* agar *user* (penguna) mudah dalam menggunakan sistem komputasi *BlackBox* yang dikembangkan.

**Kata Kunci :** Komputasi *Blackbox*, Mutu Soal, Tipe Soal Objektif

### A. PENDAHULUAN

Analisis butir soal merupakan kegiatan yang harus dilakukan guru untuk menjamin dan meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang

Makalah dipresentasikan dalam Seminar Nasional Matematika dan Pendidikan Matematika dengan tema "*Penguatan Peran Matematika dan Pendidikan Matematika untuk Indonesia yang Lebih Baik*" pada tanggal 9 November 2013 di Jurusan Pendidikan Matematika FMIPA UNY

setiap penilaian. Tujuan analisis adalah untuk mengkaji, dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan untuk memberikan penilaian kepada siswa. Khususnya untuk soal obyektif (pilihan ganda), secara empiris butir-butir soal disebut bermutu apabila memenuhi 5 (lima) persyaratan, yaitu (1) option-option pengecoh berfungsi untuk menjebak siswa yang menjawab secara tebak-tebakan atau tidak paham konsep, (2) butir soal memiliki tingkat kesulitan yang ideal, artinya tidak terlalu sulit dan tidak terlalu mudah, (3) butir soal dapat membedakan antara siswa yang menguasai konsep dengan yang tidak menguasai konsep, (4) bersifat valid, artinya tepat digunakan untuk menentukan nilai siswa dan (5) bersifat reliabel, artinya butir soal tersebut konsisten dalam memberikan interpretasi nilai bagi siswa.

Guna memperoleh soal yang bermutu, maka perlu dilakukan analisis kualitatif dan kuantitatif. Analisis butir soal secara kualitatif dilaksanakan berdasarkan kaidah penulisan soal (tes tertulis, perbuatan, dan sikap). Aspek yang diperhatikan dalam penelaahan secara kualitatif adalah setiap soal ditelaah dari segi materi, kontruksi, bahasa/budaya, dan kunci jawaban/pedoman penskoran. Sedangkan analisis soal secara kuantitatif adalah penelaahan butir soal didasarkan pada data empirik dari butir soal yang bersangkutan. Data empirik diperoleh dari soal yang telah diujikan. (Depdiknas:2008)

Pendekatan dalam analisis butir dapat dilakukan secara klasik dan modern. Analisis butir soal secara klasik adalah proses penelaahan butir soal melalui informasi dari jawaban peserta didik guna meningkatkan mutu butir soal. Kelemahan dari metode klasik ini memerlukan perhitungan yang lama dan harus disertai ketelitian yang tinggi juga kemampuan guru untuk melakukannya. Hal inilah yang menyebabkan banyak guru untuk tidak melaksanakan uji mutu soal yang diujikan. Pendekatan yang kedua menggunakan metode yang sudah modern yaitu menggunakan Item Response Theori (IRT) atau teori jawaban butir soal. Teori ini merupakan suatu teori yang menggunakan fungsi matematika untuk menghubungkan antara peluang menjawab benar suatu soal dengan kemampuan siswa. Beberapa program yang sudah tersedia untuk menganalisa mutu tes antara lain adalah ITEMAN, EXCEL, SPSS, ANATES dan lainnya. Masing-masing program ini memiliki kelebihan dan kekurangan sebagai berikut:

**Tabel 1.1 Kelebihan dan kekurangan program komputasi analisis mutu tes**

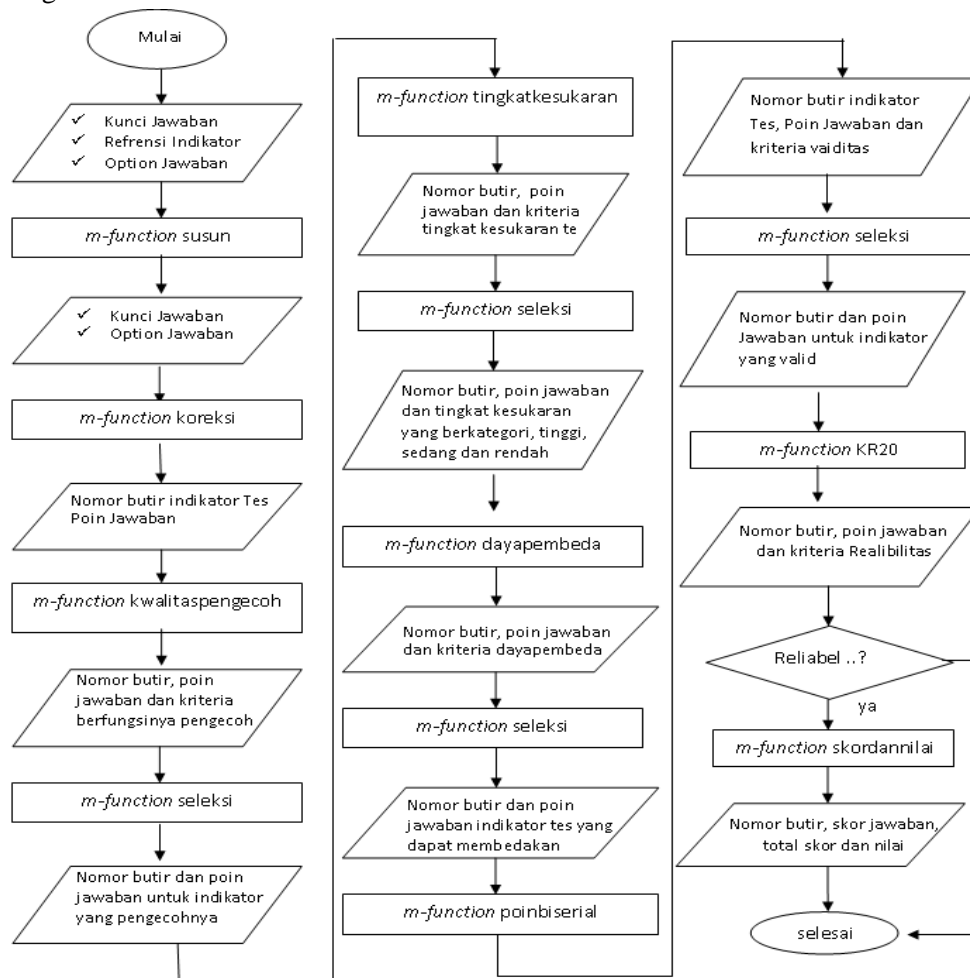
NO	PROGRAM KOMPUTASI	KELEBIHAN	KEKURANGAN
1.	ITEMAN	1. Mengoreksi jawaban siswa secara otomatis 2. Dapat melakukan analisis mutu secara simultan	1. Pengoreksian terbatas pada sebuah tipe soal 2. Tidak menampilkan proses matematis analisisnya sebagai bahan pembelajaran.
2.	ANATES	1. Mengoreksi jawaban siswa 2. Dapat melakukan analisis mutu secara simultan	1. Pengoreksian terbatas pada sebuah tipe soal 2. Tidak menampilkan proses matematis analisisnya sebagai bahan pembelajaran
3.	EXCEL	1. Penyusunan data base 2. Dapat membuat komputasi matematis	Tidak tersedia fungsi khusus untuk mengoreksi jawaban siswa dan analisis mutu, sehingga harus membuat sendiri
4.	SPSS	Fungsi matematika dan statistik yang banyak	Memerlukan kemampuan professional untuk analisis dan interpretasinya

Pada tabel di atas, program ITEMAN dan ANATES merupakan program yang secara khusus dibangun untuk menganalisis mutu butir soal. Akan tetapi keduanya tidak dapat memiliki kemampuan untuk multi tipe jenis tes dan proses matematis analisis tidak ditampilkan sehingga bagi mahasiswa atau guru yang ingin belajar meningkatkan kemampuannya sulit untuk memahami secara proses matematis analisis tersebut.

Sistem komputasi *BlackBox* dengan komponen *Input-Proses-Output* merupakan suatu sistem dimana pengguna hanya perlu memberikan inputan data jawaban siswa dengan berbagai tipe soal yang diberikan. Kemudian dengan sebuah *button* proses secara simultan oleh sistem komputasi jawaban tersebut dilakukan pengoreksian, skorsing, analisis mutu dan menampilkan output secara otomatis sesuai fungsi matematis yang dikonstruksi pada sistem. Analisis mutu mencakup pengukuran secara kuantitatif kualitas pengecoh, tingkat kesukaran, daya beda, validasi dan reliabilitas.

## B. PEMBAHASAN

Sistem komputasi *BlackBox* untuk uji mutu tes, dikembangkan dengan bahasa pemrograman *Matrix Laboratorium* (MATLAB) versi 2013. Interaksi komputasi antara item uji mutu dikembangkan berbasis *m-function*. Berdasarkan hasil kajian yang telah dilakukan, untuk efektifitas dan efisiensi, sebanyak 10 (sepuluh) *m-function* dikembangkan yaitu (1) *m-function* susun, (2) *m-function* koreksi, (3) *m-function* kualitas butir pengecoh, (4) *m-function* tingkat kesukaran, (5) *m-function* daya pembeda, (6) *m-function* poin biserial, (7) *m-function* KR20, (8) *m-function* Skor dan nilai dan (10) *m-function* uji mutu tes. Hubungan *m-function* di sajikan dalam diagram alir berikut:



Gambar 5.1 Diagram alir hubungan *m-function* untuk uji mutu tes

Hasil pengembangan *m-function* tersebut dapat dilihat pada lampiran 2 (dua), sedangkan konsep matematika dan statistik untuk komputasi dari *m-function* tersebut disajikan dalam tabel berikut:

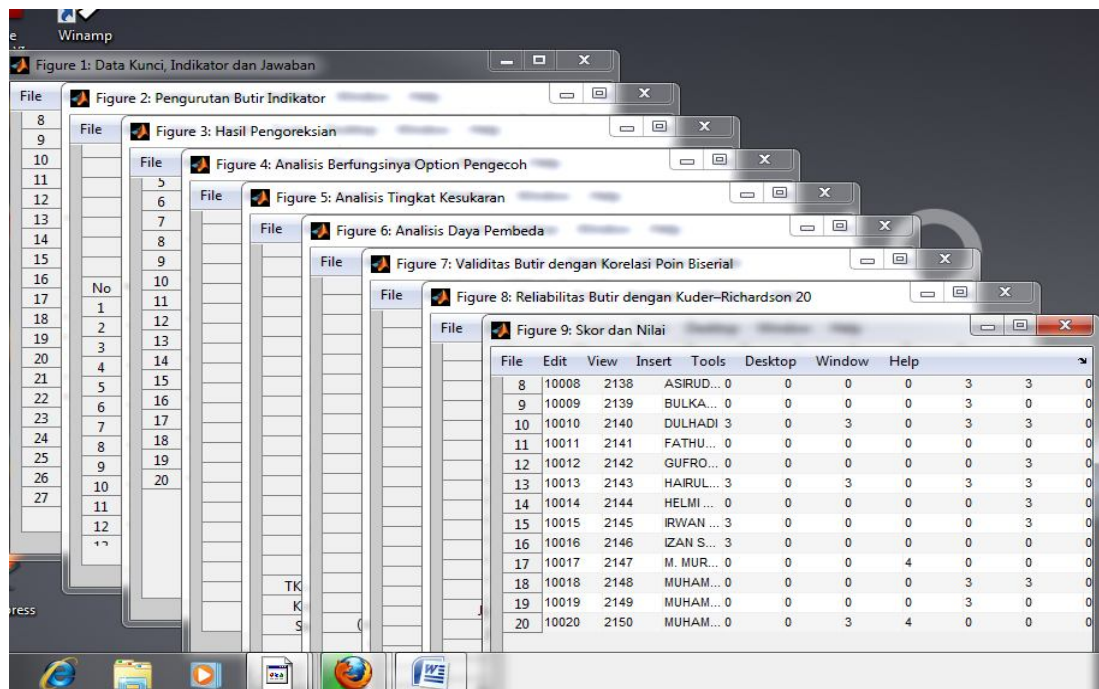
**Tabel 5.4 Matriks konten isi teori matematika *m-function* sitem komputasi *black box* untuk uji mutu tes obyektif**

No	<i>m-function</i>		Inputan	Teori Matematika	Output	
	Nama	Diskripsi			Fungsi	Laporan
1	susun	Penyusunan kunci dan butir jawaban tes berdasarkan indikator yang sejenis	1. Kunci Jawaban 2. Referensi indikator 3. Jawaban Siswa	1. Membaca banyaknya tipe soal, butir soal dan peserta tes. 2. Untuk setiap peserta tes, komputer mengidentifikasi tipe soal yang dikerjakan, kemudian mengurutkan butir jawaban siswa berdasarkan referensi nomor indikator. 3. Untuk setiap tipe soal, butir-butir kunci jawaban diurutkan berdasarkan nomor indikator	Kunci jawaban, referensi indikator dan option jawaban siswa yang telah diurutkan berdasarkan tipe dan nomor indikator yang sejenis dengan ekstensi <i>string</i>	Tabel kunci dan jawaban siswa yang telah diurutkan berdasarkan tipe dan nomor indikator yang sejenis
2	koreksi	Mengkoreksi jawaban siswa berdasarkan kunci jawaban	1. Kunci Jawaban 2. Jawaban Siswa	1. Membaca banyaknya tipe soal, butir soal dan peserta tes. 2. Untuk setiap peserta tes, komputer mengidentifikasi tipe soal yang dikerjakan, kemudian membandingkan tiap butir jawaban dengan butir kunci. Jika sama diberikan point 1 dan sebaliknya 0.	Nomor butir dan point tiap butir jawaban siswa dengan ekstensi <i>numerik</i>	Tabel identitas siswa, point tiap jawaban siswa, banyaknya benar siswa dan prosentase benar siswa dengan ekstensi <i>string</i>
3	Kwalitas pengecoh	Menentukan apakah option pengecoh berfungsi untuk menjebak siswa yang tidak menguasai konsep	1. Butir indikator dan point jawaban siswa 2. Tingkat signifikansi	Deskriptif : $P = \frac{n(salah)}{n} 100\%$ Pengecoh berfungsi jika $P \geq 5\%$	Matriks dengan isi nomor butir tes, poin jawaban dan kriteria berfungsi pengecoh (berfungsi = 1, tidak berfungsi = 0)	Tabel nomor butir, poin jawaban, besaran parameter berfungsi pengecoh (Pemilih pengecoh $\geq 5\%$ ) dan kriterianya
4	Seleksi butir	Membuat matriks baru yang memuat data butir tes yang memenuhi kriteria	1. Nomor butir dan poin jawaban 2. Kriteria uji (1 = memenuhi syarat, 0 tidak memenuhi syarat)	1. Membaca banyaknya butir tes 2. Untuk tiap butir tes, jika kriteria valid =1, maka data nomor butir dan poin jawaban disimpan pada suatu matriks baru (sebut sebagai out)	Matriks yang berisi nomor butir tes, point jawaban dan kriteria validitas (Valid = 1 dan tidak valid = 0)	Tidak ada
5	Tingkat kesukaran	Untuk mengetahui tingkat kesulitan butir soal sebagai	1. Butir indikator dan point jawaban siswa	$TK = \frac{nB}{N} 100\%$ $0\% \leq TK < 15\%$ Sangat Sukar, dibuang $15\% \leq TK < 30\%$ Sukar	Matriks nomor butir tes, poin jawaban dan kriteria	Tabel poin jawaban siswa, besaran parameter

No	m-function		Inputan	Teori Matematika	Output	
	Nama	Diskripsi			Fungsi	Laporan
			dalam bentuk 1 untuk benar dan 0 untuk salah 2. Tingkat signifikansi	$30\% \leq TK < 70\%$ Sedang $70\% \leq TK < 85\%$ Mudah $85\% \leq TK \leq 100\%$ , Sangat Mudah, dibuang	tingkat kesukaran (3=sulit, 2=sedang dan 1=mudah)	tingkat kesukaran dan kriteria tingkat kesukaran (Tinggi, sedang dan mudah)
6	Daya pembeda	Untuk mengetahui apakah butir soal dapat membedakan siswa yang menguasai konsep dan tidak menguasai konsep	1. Butir indikator dan point jawaban siswa dalam bentuk 1 untuk benar dan 0 untuk salah 2. Tingkat signifikansi	$DP = \frac{BA - BB}{NA} 100\%$ $DP \geq 50\%$ Sangat Baik $30\% \leq DP < 50\%$ Baik $20\% \leq DP < 30\%$ Agak Baik, Revisi $10\% \leq DP < 20\%$ Buruk, dibuang $DP < 10\%$ Sangat Buruk, diibuang	Matriks nomor butir tes, poin jawaban dan kriteria daya beda (dapat membedakan = 1, dan tidak membedakan = 0)	Tabel nomor butir, poin jawaban, besaran parameter daya pembeda dan kriteria daya beda
7	Poin biserial	Mengetahui apakah butir indikator tes valid/tepat digunakan untuk mengukur hasil belajar siswa	1. Butir indikator dan point jawaban siswa dalam bentuk 1 untuk benar dan 0 untuk salah 2. Tingkat signifikansi	$r_{pb} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$ Jika $r_{pb} \geq r_{inv(n,5\%)}$ maka butir tes valid	Nomor Butir soal dan tingkat	Tabel poin jawaban peserta tes, besaran parameter poin biserial dan kriteria validitas
8	Kr20	Menentukan apakah butir-butir tes reliabel/konsisten untuk mengukur hasil belajar siswa	1. Butir indikator dan point jawaban siswa dalam bentuk 1 untuk benar dan 0 untuk salah 2. Tingkat signifikansi	$r_h = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum pq}{s_{total}^2} \right)$ Jika $r_{pb} \geq r_{inv(n,5\%)}$ maka butir tes reliabel	Matriks yang memuat butir tes, poin jawaban dan kriteria reliabilitas (reliabilitas = 1, tidak reliabilitas = 0)	Tabel poin jawaban peserta tes, besaran parameter KR20 dan kriteria reliabilitas
9	Skor dan penilaian	Untuk menghitung skor dan nilai dari tiap peserta tes	1. Butir indikator dan point jawaban siswa dalam bentuk 1 untuk benar dan 0 untuk	Butir tes yang sulit diberikan skor 3, sedang = 2 dan mudah = 1 $Nilai = \frac{\text{Jumlah skor}}{\text{Skor ideal}} \times 100$	Matriks nomor butir tes, poin jawaban dan kriteria tingkat kesukaran (3=sulit, 2=sedang dan 1=mudah)	Tabel skor jawaban siswa, total skor dan nilai

No	<i>m-function</i>		Inputan	Teori Matematika	Output	
	Nama	Diskripsi			Fungsi	Laporan
			salah 2. Tingkat signifi- kansi			
10	Uji mutu tes	Untuk melakukan komputasi uji mutu tes secara simultan	Data numerik butir indikator dan point jawaban siswa Kriteria tingkat kesukaran	Menjalankan komputasi sebagaimana diagram alir pada gambar 5.2.	Tidak ada	9 jenis tabel mencakup 9 <i>m-function</i> di samping.

Setelah komputasi menggunakan *m-function* dikembangkan/dibuat lalu dilakukan simulasi untuk mengetahui kesesuaian antara teori matematika analisis mutu tes dengan output komputasi. Output komputasi yang diperoleh berupa 9 (Sembilan) tabel hasil data dan analisis meliputi (1) data kunci, indikator dan jawaban, (2) Pengurutan butir indikator, (3) Hasil Pengoreksian, (4) Analisis berfungsinya option pengecoh, (5) Analisis tingkat kesukaran, (6) Analisis daya pembeda, (7) Validitas butir dengan korelasi poin biserial, (8) Reliabilitas butir dengan korelasi KR20 dan (9) skor dan nilai.



Gambar 5.1 Output *m-function* ujimututes

Tabel ke-1 merupakan rekaman data, referensi indikator dan option jawaban siswa, tabel ke-2 data hasil pengurutan butir tes berdasarkan kesamaan indikator, tabel ke-3 rekaman hasil pengoreksian jawaban siswa, tabel ke-4 rekaman analisis berfungsinya option pengecoh, tabel ke-5 rekaman hasil analisis tingkat kesukaran, tabel ke-6 rekaman hasil analisis tingkat daya pembeda butir tes, tabel ke-7 rekaman hasil analisis validitas butir, tabel ke-8 rekaman hasil uji reliabilitas butir dan tabel ke-9 rekaman hasil analisis skor dan penilaian. *Function* ini merupakan *function* komputasi yang akan digunakan dalam pengembangan aplikasi *project system* komputasi *BlackBox* yang berfungsi untuk melakukan analisis mutu butir tes secara simultan (otomatis).

Dari uraian di atas, maka pengembangan *m-function* sebagai komponen dari system komputasi *BlackBox* telah dapat dikembangkan dengan kinerja komputasi yang valid.



### C. KESIMPULAN

Berdasarkan hasil penelitian yang telah dicapai, maka pengembangan sistem komputasi *BlackBoox* untuk uji mutu dapat dilakukan dengan langkah sebagai berikut :

1. Untuk efektif dan efisiensi pengguna, maka inputan sistem komputasi dapat dilakukan di *excel* dengan jenis data masukan *string* mencakup data kunci jawaban, referensi butir indikator dan option jawaban siswa.
2. Untuk efektif dan efisien, maka proses komputasi dibangun dalam bentuk *m-function* mencakup (1). *m-function* susun, (2) *m-function* koreksi, (3) *m-function* kualitas pengecoh, (4) *m-function* tingkat kesukaran, (5) *m-function* daya pembeda, (6) *m-function* poin biserial, (7) *m-function* KR20, (8) *m-function* skor dan nilai, (9) *m-function* seleksi dan (10) *m-function* uji mutu tes.
3. Untuk efektifitas dan efisiensi pelaporan dan proses pembelajaran, maka output komputasi disajikan dalam tabel mencakup sumber data analisis, deskripsi parameter data dan konklusi dari hasil analisis

### DAFTAR PUSTAKA

- Aiken, Lewis R. (1994). *Psychological Testing and Assessment*, (Eight Edition), Boston: Allyn and Bacon.
- Ali. M.,(2012), *Analisis Butir Soal Dengan ANBUSO*, Makalah Pengabdian Masyarakat, FE UNY, Yogyakarta.
- Anwar., (2012), *Analisis Butir Soal Pilihan Ganda sebagai Bahan Rujukan Bagi Guru*, Makalah seminar
- Bejar, Issac I. (1983). *Introduction to Item Response Theory and Their Assumptions*. Hambleton, Ronald K (Editor). *Appliction of Item Renponse Theory*. Canada: Educational Research Institute of British Columbia.
- Bloom, Benyamin S., dkk (1971), *Handbook on Formative and Summative Evaluation of Student Learning*, McGraw-Hill Book Company, New York.
- Departemen Pendidikan dan Kebudayaan RI (1989), *Pedoman Penelaahan, Perbaikan dan Perakitan Soal*, Jakarta.
- Depdiknas. (2008). *Panduan Analisis Butir Soal*. Jakarta: Departemen Pendidikan Nasional
- Gronlund, Norman E. (1985) *Measurement and Evaluation in Teaching, Fifth Edition*, Macmillan Publishing Company New York.
- Guilford (1973). *Fundamental Statistic in Psychology and Education*, Tokyo: Mc Graw-Hill Kogakusha.
- Irmayansyah. (2011), Pengaruh Penggunaan Program Aplikasi Anates Versi 4 Terhadap Kinerja Guru Dalam Menganalisis Butir Soal, *Jurnal Ilmiah Teknologi dan Informasi*, Volume 2
- ITEMAN., (2006). *User's Manual for the ITEMAN Conventional Item Analisis Program Second Edition*, Minnesota: Assasment System Corporation.
- Karno To (1995). *ANATES Versi 2.5, Program Komputer khusus untuk analisis tes obyektif dan uraian*. Bandung: FIP IKIP Bandung.
- Moleong, Lexy J., (2002). *Metodologi Penelitian Kualitatif*. Remaja Rosdakarya, Bandung.
- Murphy, Kevin R. & Charles O. Davidshofer (1988). *Psychological Testing, Principles & Applications*. New Jersey: Prentice Hall International Inc.

- 
- Nitko, Anthony J. (1996), *Educational Assessment of Student, 2nd Edition*, Prentice Hall, New Jersey, Ohio.
- Subino (1987), *Konstruksi dan Analisis Tes, Suatu Pengantar Kepada Teori Tes dan Pengukuran*, Proyek LPTK Depdikbud, Jakarta.
- Suharsimi Arikunto., (2003). *Dasar-dasar Evaluasi Pendidikan*. Bumi Aksara. Jakarta
- Widodo., (2006). Taksonomi Bloom dan Pengembangan Butir soal. *Buletin Puspendik*